24-2594

IN THE

United States Court of Appeals

FOR THE SECOND CIRCUIT

MICHELLE PHIPPEN, INDIVIDUALLY AND AS GENERAL GUARDIAN OF P.P. AND L.A., MINORS, ALISHA DAY, INDIVIDUALLY AND AS MOTHER, GENERAL GUARDIAN OF A.D., A MINOR, SARAH STOKES, INDIVIDUALLY AND AS GENERAL GUARDIAN OF K.G., A MINOR, JUAN EMANUEL BORDOY, INDIVIDUALLY, MARY ELIN ARCE, INDIVIDUALLY AND AS MOTHER OF JUAN EMANUEL BORDOY, AMANDA TRIGLOFF, INDIVIDUALLY AND AS GENERAL GUARDIAN OF R.S., A MINOR, DEANDRE BARBEE, INDIVIDUALLY, JANTAIL BARBEE, INDIVIDUALLY AND AS MOTHER OF DEANDRE BARBEE, LAURIE COURINGTON, HUNTER COURINGTON, JENNIFER MORROW, ALENA MORROW, CALLISTA BASSETT,

(Caption continued on inside cover)

ON APPEAL FROM THE UNITED STATES DISTRICT COURT
FOR THE SOUTHERN DISTRICT OF NEW YORK

REPLY BRIEF FOR PLAINTIFFS-APPELLANTS [REDACTED]

ASHLEY C. KELLER
ASHLEY L.F. BARRIERE
KELLER POSTMAN LLC
150 N. Riverside Plaza
Suite 4100
Chicago, IL 60606
(312) 741-5222

JOHN J. SNIDOW
JOHN M. MASSLON II
ROSEANN R. ROMANO
KELLER POSTMAN LLC
1101 Connecticut Avenue NW
Suite 1100
Washington, DC 20036
(202) 378-0649

Attorneys for Plaintiffs-Appellants Sonnita Roby, Taylor Brown, Michelle Phippen, Amanda Trigloff, Laurie Courington, Hunter Courington, Jennifer Morrow, Alena Morrow, Callista Bassett, Andrew Bassett, Shannon Mikuski, Braydon McKenzie, Heather Gilliam, Colby Gilliam, and Michelle Brown

(*Counsel continued on inside cover*)

ANDREW BASSETT, SONNITA ROBY, INDIVIDUALLY AND AS GENERAL GUARDIAN OF D.P., A MINOR, SHANNON MIKUSKI, BRAYDON MCKENZIE, HEATHER GILLIAM, COLBY GILLIAM, TAYLOR BROWN, TARYNE BURKE, INDIVIDUALLY AND AS MOTHER WITH COURT APPOINTED GUARDIAN OF ASHTON BURKE, SAMARI SIMS, INDIVIDUALLY, DENISA CULLOM, INDIVIDUALLY AND AS MOTHER OF SAMARI SIMS, COLLIN STOVER, INDIVIDUALLY, DANA STEWART, INDIVIDUALLY AND AS MOTHER OF COLLING STOVER, RIAN CZAR JOHNSON, INDIVIDUALLY, SHILO RICH, INDIVIDUALLY AND AS MOTHER OF RIAN CZAR JOHNSON, SHEENA SCHNEPP, INDIVIDUALLY AND AS MOTHER AND NATURAL GUARDIAN OF H.S., A MINOR, ZAYNE COSTELLO, INDIVIDUALLY, CRYSTAL ALEXANDER, COURTNEY TILLOTSON, MICHELLE BROWN,

Plaintiffs-Appellants,

—against—

WALGREENS CO., JOHNSON & JOHNSON CONSUMER INC., WALMART INC.,

Defendants-Appellees.

DANIEL C. BURKE
BERNSTEIN LIEBHARD LLP
10 East 40th Street, 22nd Floor
New York, New York 10016
(212) 779-1414

Attorneys for Plaintiffs-Appellants Juan Emanuel Bordoy, Mary Elin Arce, Deandre Barbee, Jantail Barbee, Taryne Burke, Samari Sims, Denisa Cullom, Collin Stover, Dana Stewart, Rian Czar Johnson, Shilo Rich, Sheena Schnepp, Zayne Costello, Crystal Alexander, and Courtney Tillotson

LINDSEY SCARCELLO
WAGSTAFF & CARTMELL, LLP
4740 Grand Avenue, Suite 300
Kansas City, Missouri 64112
(816) 701-1100
lscarcello@wcllp.com

Attorneys for Plaintiff-Appellant Alisha Day

TABLE OF CONTENTS

Page

TABLE OF AUTHORITIES

Page(s)

Cases

Other Authorities

PRELIMINARY STATEMENT

Defendants never grapple with a straightforward issue of law: an expert's method is reliable, and reliably applied, if it follows what actual scientists say and do in the field. *See* Appeal No. 916 Reply 1–3; Appeal No. 916 Opening Br. 28–29; *cf. Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 152 (1999). Incanting the words "abuse of discretion," Defs. Br. at 29, is not a license authorizing district court judges to determine the "state of the science," Defs. Br. at 2, let alone for them to declare that the science can yield only one conclusion that *independent*, *published* experts disagree with. Appropriately cabining discretion so it fits the skillset of jurists is the only way to ensure that lay judges do not "assess the weight of conflicting evidence," *In re Joint E. & S. Dist. Asbestos Litig.*, 52 F.3d 1124, 1131 (2d Cir. 1995) (citation omitted), or play "amateur scientist," *id.* at 1137.

Respecting the limited gatekeeping role contemplated by Rule 702 does not gut the abuse-of-discretion standard. District courts have wide latitude in the first instance to hear and evaluate expert testimony to determine that it is grounded in reliable science. *Ruiz-Troche v. Pepsi Cola of Puerto Rico Bottling Co.*, 161 F.3d 77, 81 (1st Cir. 1998). They may flexibly determine if unreliable methods that infect part of an expert's opinion are serious enough to warrant wholesale exclusion. *Cf. Amorgianos v. Nat'l R.R. Passenger Corp.*, 303 F.3d 256, 264 (2d Cir. 2002). And they retain the discretion to survey the independent scientific literature—*without*

needing to become experts in the underlying science—to assess if the expert's bottom line is supported by more than mere *ipse dixit*. *Gen. Elec. Co. v. Joiner*, 522 U.S. 136, 146 (1997). None of that saves the District Court's opinion. The District Court never saw Dr. Ness testify, never identified any unreliable methods, dismissed Dr. Ness's extensive assessment of contrary evidence, and decided for itself that the "state of the science" does not allow Dr. Ness to fill gaps that *independent scientists* themselves have filled. SPA-256. That approach aims not to assess the reliability of expert testimony but to determine if it is correct. It converts district courts from gatekeepers to "the role of St. Peter at the gates of heaven," an anointment that does not accompany an Article III commission. *McCullock v. H.B. Fuller Co.*, 61 F.3d 1038, 1045 (2d Cir. 1995). The judgment below must be reversed.

ARGUMENT

I. The District Court Erred By Excluding Dr. Ness.

Each step of Dr. Ness's analysis was reliable. She reliably considered genetic confounding and each Bradford Hill factor. Independent scientists echo her approach *and* her conclusions. *See, e.g.*, 916.A-4467 (Alemany (2021)); 916.A-4488 (Bauer & Kriebel: "Several lines of reasoning suggest that bias, confounding and chance are not solely responsible."). There is no room for lawyers or judges to second guess those well-grounded, even if debatable, scientific judgments.

A. Dr. Ness Reliably Considered And Ruled Out Genetic Confounding.

Given the District Court's first Rule 702 decision, Dr. Ness painstakingly evaluated the possibility that the repeatedly observed association between prenatal APAP exposure and ADHD in offspring is somehow due to genetic confounding. A-7422–27.[1] For genetic confounding to explain the study results, the genes that cause a child to develop ADHD must also spur women to take more APAP while pregnant (but not before or after pregnancy). Is it possible that some as-yet undiscovered genes in women both cause ADHD in their children and also cause more APAP-treating headaches, minor aches, pains, and fevers *only* during pregnancy? It, like anything in science, is possible. But Dr. Ness considered all of the literature that could possibly bear on the question at the time of her analysis to reach the reliable conclusion that genetic confounding was unlikely to explain the entire association. There is no yawning scientific gap between the evidence Dr. Ness reliably analyzed and the conclusion she reached. *Cf. Joiner*, 522 U.S. at 146. The Court can be perfectly sure of that point because scores of independent scientists have filled the same gaps to reach Dr. Ness's conclusion. *See, e.g.*, 916.A-1389 (Liew et al. noting genetics "do not explain the association observed for acetaminophen exposure");

[1] Dr. Ness devoted a separate section of her report to examining genetic confounding, A-7422–27, but also considered it throughout her literature review, *see, e.g.* A-7292–94 (discussing at length the treatment of genetic confounding in Ystrom (2017)).

916.A-4535 (Stergiakouli et al. stating association "is not explained by unmeasured familial factors," *i.e.*, genetics); 916.A-4762 (Bornehag et al.: "[C]onfounding alone is an unlikely explanation for the associations reported in these studies."); 916.A-4468 (Alemany et al. concluding the same); 916.A-4560 (Gou et al. concluding the same).

Defendants attempt to brush these independent scientists aside by declaring, albeit in a footnote, that their analysis is of "little relevance" because they predate Leppert (2019), Gustavson (2021), and Ahlqvist (2024). Defs. Br. at 34 n.11. Appellate attorneys are hardly qualified to say whether the opinions of dozens of independent scientists are now of "little relevance" based on later literature that Defendants happen to like. Take it from the Gustavson authors themselves, who recently cautioned that sibling-control studies "may sometimes introduce more bias than they alleviate." Kristin Gustavson et al., *Familial Confounding or Measurement Error? How to Interpret Findings from Sibling and Co-Twin Control Studies*, 39 Eur. J. Epidemiology 587, 587 (2024). And nothing in those studies undermines the robust negative controls that prior studies deployed to control for genetics. A-7422, A-7425. That is why independent scientists to this day continue saying what Dr. Ness concluded in her report. John P. Jones III et al., *Evaluating the Role of Susceptibility Inducing Cofactors and of Acetaminophen in the Etiology of Autism Spectrum Disorder*, 14 Life 1, 1 (2024); Brennan Baker et al., *Associations*

of Maternal Blood Biomarkers of Prenatal APAP Exposure with Placental Gene Expression and Child Attention Deficit Hyperactivity Disorder, 3 Nature Mental Health 318, 318, 324–25 (2025) (finding APAP detection in maternal plasma triples risk of ADHD in offspring, and identifying similar altered placental expression in human and animal studies due to APAP that would undercut maternal genetics as explanation for association).

The closest Defendants come to identifying a methodological flaw in Dr. Ness's analysis is that she "cherry picked" the literature as part of her "result-driven" analysis. Defs. Br. at 32–33. To be sure, a scientist who undertakes an analysis to reach a preordained conclusion may be unreliable. That is vastly different from a scientist who considers all of the available evidence and concludes, on balance, that it supports a position Defendants (and the District Court) disagree with. That is what happened here. Dr. Ness extensively considered both Brandlistuen (2013) and Gustavson (2021). A-7342–45; A-7318–23. Defendants do not engage with most of Dr. Ness's analysis, including her mathematical proof that genetic confounding cannot explain the entire association.[2] A-7422; 916.A-4455; Pls. Br. at 30; 34–35.

[2] Although the District Court accused Dr. Ness of "prioritize[ing]" Brandlistuen (2013) over Gustavson (2019), SPA-235, Dr. Ness's report says otherwise. She repeatedly emphasized—in response to the District Court's directive—her focus on studies with diagnostic endpoints and did not ever explicitly state that she placed more weight on Brandlistuen (2013) than Gustavson (2019). A-7342 (discussing Brandlistuen as one of "other studies . . . that are important for context"). She

They simply decree that Dr. Ness should have given more weight to Gustavson, which is the study with results they prefer.

But as Dr. Ness thoroughly explained, Gustavson was severely underpowered with only 34 doubly discordant pairs, a limitation the authors themselves shared. A-7318–23; Kristin Gustavson et al., *Supporting Information for Acetaminophen Use During Pregnancy and Offspring Attention Deficit Hyperactivity Disorder – a Longitudinal Sibling Control Study* at 7, https://perma.cc/9JYW-UEKG (Given the small sample size, "statistical power to detect within effects was relatively low. Hence, these results should be interpreted with caution."). The Brandlistuen sibling study, by contrast, contained 939 doubly discordant pairs, providing much more confidence in its findings. 916.A-1244. It is fair enough that Gustavson focused on confirmed cases of ADHD, while Brandlistuen examined symptoms of the condition using a reliable and validated instrument. But all studies have strengths and weaknesses. It is not "cherry picking" or "result-driven" for an eminently qualified scientist to *explain* why she believes one study—limitations and all—better captures what is likely to be true than another study that, in context, has more serious weaknesses. *See* A-7318–23, A-7424–25. If Dr. Ness was unreliable on this score, it can only be because this Court holds that the District Court has discretion to pick

---

certainly criticized Gustavson (2019), but there is no evidence of this sort of "result-driven" analysis.

between study strengths and weaknesses. That cannot be the law. *See Elosu v. Middlefork Ranch Inc*., 26 F.4th 1017, 1026 (9th Cir. 2022) ("Rule 702 does not license a court to engage in freeform factfinding, to select between competing versions of the evidence, or to determine the veracity of the expert's conclusions.").

Turning from Dr. Ness's actual analysis, Defendants remarkably champion the District Court's extensive, three-page review of evidence *Dr. Ness could not have analyzed*: Ahlqvist (2024). As should be plain, it cannot be an unreliable application of methods and principles to fail to assess evidence that did not exist. Pls. Br. at 21–22; 35–36. Begrudgingly acknowledging the point, Defendants first observe that the District Court claimed it would have reached the same exclusion decision without the study. Defs. Br. at 36. So apparently, no harm, no foul. That begs a simple question: why did the District Court spill so much ink considering complicated scientific evidence—without *any* expert analysis from *either* side—if it had no bearing on the Rule 702 inquiry? The answer is obvious. The District Court was not about to pass up an opportunity to support its side of a scientific debate. The District Court said as much expressly. SPA-237 ("The Court will not ignore this study.").

After asserting in one breath that Ahlqvist was irrelevant to the District Court, Defendants' next breath claims it was proper to consider because Dr. Ness could have talked about the study at her deposition and it was raised in the parties' briefs.

Defs. Br. at 35. The Court should not be fooled by Defendants' distortion of the record. The Ahlqvist study was published *the day of* Dr. Ness's deposition. *See* A-8915 (Tr. 386:4–7) ("And I'll represent to you that [Ahlqvist (2024)] is an article that just became available today, perhaps within – within hours."). How Defendants knew the study was being released and obtained a copy "perhaps within hours" of when it was first available is a question for another day. It is ludicrous to claim that Dr. Ness had an opportunity to express an expert opinion about a study that defense counsel laid before her for the first time while she was under oath.[3] The parties' Rule 702 briefing was similarly no excuse for the District Court to use Ahlqvist to criticize Dr. Ness's methods and principles. Plaintiffs mentioned Ahlqvist in their Rule 702 brief to make that very point: a study that came out *after* an expert's analysis cannot bear on reliability. MDL Dkt. 1477 at 18–20. On Defendants' telling, litigation briefs opposing consideration of a study somehow authorized the District Court to weigh it. There is simply no defense of the District Court's approach, and no denying that it reveals a judge determined to resolve scientific questions for herself.

[3] This is nothing like *In re Zoloft (Sertraline Hydrochloride) Prods. Liab. Litig.*, No. 12-MD-2342, 2015 WL 7776911, at *7 (E.D. Pa. Dec. 2, 2015), Defs. Br. at 35, where the study was published before the expert's deposition and he testified that he was "familiar" with it.

Rule 702 does not permit this for good reason. Since its publication, numerous independent scientists have criticized Ahlqvist for its significant weaknesses. *See* Pls. Br. at 35, n.3. Indeed, *after* Plaintiffs' opening brief, still more independent scientists have criticized the study. *See* Baker et al., *supra*, at 325 (noting the significant under-exposure of 7.5% likely results in exposure misclassification, "the most likely consequence of [which] is bias toward the null, and this type of bias more strongly impacts 'within-pair' estimate from sibling control analyses compared with the overall population estimate."). Relying on Baker, the new head of the Centers for Medicare and Medicaid Services has urged pregnant women to find safer, non-medicinal ways to control discomfort. Mehmet Oz & Michael Roizen, *Reduce Your Child's Risk of ADHD*, Middletown Press (Apr. 7, 2025), https://www.pressreader.com/usa/the-middletown-press-middletown-ct/20250407/281831469556862. Without the benefit of any expert opinion, the District Court identified not a single weakness in Ahlqvist and touted it as support for excluding Dr. Ness. The District Court's decision to assume the role of "amateur scientist" must be reversed. *In re Joint E. & S. Dist. Asbestos Litig.*, 52 F.3d at 1137.

B. Dr. Ness Reliably Applied The Bradford Hill Factors.

1. Consistency

Dr. Ness reliably evaluated consistency. A-7381–85; Pls. Br. at 37–38. Other independent scientists agree. *See, e.g.*, 916.A-4488 (Bauer & Kriebel: "First, there

is the consistency; all nine studies suggest a moderate increase in risk."); 916.A-4474 (Olsen & Liew: "[F]ive different prospective cohorts have consistently estimated a positive link."); 916.A-4551 (Gou authors reaching a similar conclusion). Indeed, the Alemany authors explicitly found that the Bradford Hill factor of consistency was satisfied. 916.A-4467 ("Consistency is supported because we observed consistent results using a variety of populations and methods.").

Defendants and the District Court accuse Dr. Ness of "substantially exaggerat[ing] [the studies'] consistency" because not all trimester results were statistically significant. Defs. Br. at 37. But that is moving the goal posts. The relevant question is whether taking APAP while pregnant is associated with ADHD in offspring. It is. Stratifying results by trimester reduces sample size and increases random error, which can "make finding true subgroup effects difficult," so the "overall trial result is usually a better guide to the direction of effect in subgroups than the apparent effect observed within a subgroup." Salim Yusuf et al., *Analysis and Interpretation of Treatment Effects in Subgroups of Patients in Randomized Clinical Trials*, 266 JAMA 93, 93 (1991); *see also* Timothy L. Lash et al., *Modern Epidemiology* 751 (4th ed. 2021) ("A common misinterpretation of significance tests is to claim that there is no difference between two observed groups because the null test is not statistically significant.").

Like the District Court, Defendants next conflate heterogeneity and inconsistency, revealing once again the dangers of permitting a layperson's intuition to define the scope of reliable scientific practice. Defs. Br. at 36–37; SPA-241. To an untrained lawyer, heterogeneity and inconsistency may be "two plain-language synonyms," Defs. Br. at 36 n.12, such that heterogeneity precludes a finding of consistency. But to practicing epidemiologists, "consistency" is a Bradford Hill factor that is satisfied when an association has "been repeatedly observed by different persons, in different places, circumstances, and times[.]" 916.A-4683. It does not matter that the repeated observations of a link cover *heterogenous* populations or result from varied study designs.[4] That is what heterogeneity measures. It is a term of art used in meta-analyses[5] to describe differences amongst the pooled studies in terms of, *inter alia*, participant characteristics, study design, and variabilities in how the intervention is captured. *Cochrane Handbook for Systematic Reviews of Interventions* § 10.10.1 (Julian Higgins & James Thomas eds., 2024). That there may be heterogeneity amongst the pooled studies in Ricci (2023) or Masarwa (2020), therefore, does not inform whether the overall studies satisfy the

[4]Modern incantations affirm this definition. *See, e.g.* A-8283 (Fedak 2015) ("when multiple epidemiologic studies using a variety of locations, populations, and methods show a consistent association between two variables").

[5]*See generally* Julian P.T. Higgins & Simon G. Thompson, *Quantifying Heterogeneity in a Meta-Analysis*, 21 Stats. Med. 1539 (2002).

consistency factor of Bradford Hill. *See* SPA-241–42.[6] In fact, that so many different studies, cohorts, and variabilities (heterogeneity) *all* showed an association is *stronger* evidence that the consistency factor of Bradford Hill is met.

The District Court dismissed Dr. Ness's consistency analysis with a word game, latching onto a lawyer's intuition that when Ricci described heterogeneity across study designs it somehow meant Dr. Ness could not find the consistency factor of Bradford Hill satisfied. There is a reason no independent scientist has ever embraced Defendants' or the District Court's treatment of inconsistency and heterogeneity as "plain-language synonyms." And there is a reason Rule 702 does not deputize district courts (or defense counsel) to write their own scientific thesaurus.

2. Temporality

Dr. Ness reliably concluded that temporality was satisfied because almost all of the epidemiological studies consisted of prospective cohorts. A-7408; Fed. Jud. Ctr., *Reference Manual on Scientific Evidence* 558 (3d ed. 2011) ("One advantage of the cohort study design is that the temporal relationship between exposure and

[6] The district court cases cited by Defendants, Defs. Br. at 36 n.12, do not stand for the unscientific proposition that inconsistency and heterogeneity are synonyms that can be used interchangeably. Even if they did, this Court should not accept ill-considered, drive-by dicta from unbinding judicial opinions in the face of clear-cut scientific authority revealing the fatal flaw in the District Court's conflation of distinct concepts.

disease can often be established more readily than in other study designs, especially a case-control design."). She is not alone. Appeal No. 916 Opening Br. 38–39. The Alemany authors state that the "causal criteria supported by the current findings include . . . temporality," and go on to specify that "[t]emporality is supported because the exposure precedes the onset of the symptoms assessed." 916.A-4467.

Defendants shift from 702(d) to 702(a), questioning Dr. Ness's qualifications to opine on temporality. That is a remarkable contention. Dr. Ness is a medical doctor, and even the District Court recognized that she "is an esteemed epidemiologist." SPA-224. Bradford Hill is the seminal method by which "esteemed epidemiologist[s]" assess causation, and temporality is the only Bradford Hill factor that "epidemiologists universally agree is essential to causal inference." A-8284. So on Defendants' telling, an M.D. and a decorated epidemiologist lacks the qualifications to perform the standard method of her scientific discipline, particularly on the one factor that method demands to draw a causal inference. Defs. Br. at 39 (quoting SPA-225). No one but Defendants' lawyers think that is true. There is certainly no support in the scientific literature for the proposition that someone with Dr. Ness's credentials cannot evaluate temporality.

3. Dose Response

The District Court made up a rule: dose-response requires precise dose measurements. Scientists outside of court looking at this very association do not

practice the District Court's concocted rule of science, just as Dr. Ness did not follow it. *See* 916.A-4509 (Liew (2014): "higher use frequency increas[ed] risk in an exposure-response manner"); 916.A-4457 (Ricci (2023) finding results "suggest[ed] a dose-response effect"); 916.A-1199 (Baker (2020): "A dose-response association was detected."); 916.A-4527 (Ji (2020): "[T]here were dose-response patterns."). Notably, the Alemany meta-analysis said that the dose-response factor of Bradford Hill was satisfied. 916.A-4467.

It is not just this literature. As Dr. Ness pointed out, the CDC used *days* at Camp Lejeune, not the precise amount of water exposure, to evaluate the dose response for water contamination on base. A-7407–08; Pls. Br. at 16; 31. The District Court and Defendants cast this example aside, concluding—without any citation—that "anyone who lived or worked at Camp Lejeune would most likely have been exposed daily to the drinking water, so duration of exposure was a reasonable proxy for dosage there." SPA-251–52. Even as a matter of pure intuition, the response does not withstand scrutiny. Some people bathe twice a day, others twice a week. Some drink eight glasses of water per day, others hydrate less. When doses cannot be measured precisely, scientists are content to use days as a proxy that can still accurately capture a dose response.[7] And the exact same logic applies here.

[7] Other examples abound. For decades, epidemiologists have used "pack years" to measure the risks of smoking rather than precise quantities of cigarette carcinogens.

True, some women who take APAP on a day may consume only one pill, while others consume six. But over an entire population of women, days of exposure is a perfectly appropriate proxy for dose. Independent scientists, alongside Dr. Ness, have reliably said so. The District Court was not qualified to decree, without citation, when duration of exposure "is a reasonable proxy for dosage." SPA-251–52. That is a scientific judgment *that actual scientists have made*, not an exercise of judicial discretion.

There was similarly nothing "result-driven" about Dr. Ness's judgment that so many independent scientists are right. Just the opposite. It was the District Court and now the Defendants who proceed in a "result-driven manner." Defs. Br. at 41. The District Court criticized Dr. Ness for not acknowledging the Baker authors' statement that "they did not correlate maternal acetaminophen use with the acetaminophen concentrations in meconium." SPA-252. Talk about burying the lede. The Baker authors *concluded* that "[a] dose-response association was detected," 916.A-1199, exactly as Dr. Ness said. Ignoring that bottom-line conclusion served the District Court's "result-driven" Rule 702 analysis. It had nothing to do with the reliability of Dr. Ness's principles and methods.

See generally Roy A. Pleasants et al., *Both Duration and Pack-Years of Tobacco Smoking Should Be Used for Clinical Practice and Research*, 17 Annals ATS 804 (2020).

4. Biological Plausibility

Dr. Ness reliably opined that oxidative stress is a plausible mechanism to explain APAP's in utero neurotoxic effects. A-7387–99. She is in good company, as multiple independent scientists—[redacted]—have said the same thing. 916.A-4467 (Alemany (2021): "oxidative stress due to inflammation-induced immune activation" satisfies "biological plausibility" of Bradford Hill). Gustavson (2021) acknowledged that "[b]iologically plausible mechanisms for effect on fetal neurodevelopment include oxidative stress and neurotoxicity." 916.A-4541 (citing Ghanizadeh (2012)). The Chen (2019) authors also noted several "reported" plausible mechanisms, including that "acetaminophen may alter the intrauterine immune system and increase the predisposition for oxidative stress and inflammation, which disrupts the normal development of microglia and their interaction with neurons." 916.A-1282. [redacted] A-9401, A-9403.

The District Court *ignored* these independent authors, [redacted], and concluded that Dr. Ness failed—for the same reasons as Plaintiffs' prior experts—to "reliably fill critical gaps in the purported mechanistic pathway." SPA-255. But the gap between a method and a conclusion cannot be too big to fill if *independent scientists* have filled it in the *same way* as the testifying expert. *Cf.*

Joiner, 522 U.S. at 146. The District Court may wish that biological plausibility requires identifying the "precise physiological process or processes by which these conditions . . . develop." SPA 91–92. But science is the domain of scientists, and there is no getting around that the District Court applied a standard that the scientific community does not. *See* Appeal No. 916 Opening Br. 42–43.

Defendants retreat once again to Rule 702(a), contending that Dr. Ness is somehow unqualified to opine on biological plausibility. Defs. Br. at 42–43. Not even the District Court endorsed that argument. It is ludicrous. As noted above, Bradford Hill is the "canonical approach" to causation for epidemiologists, Dr. Ness is an M.D. and an epidemiologist, and so Dr. Ness is eminently qualified to assess all the Bradford Hill factors, including biological plausibility. 916.A-4646. Dr. Ness's biological plausibility "opinions are expressed as part of h[er] epidemiologic analysis of causation, which, as previously discussed, [Dr. Ness] is sufficiently qualified to perform." *Deutsch v. Novartis Pharms. Corp.,* 768 F. Supp. 2d 420, 459 (E.D.N.Y. 2011) (holding epidemiologist qualified to analyze biological plausibility as part of a Bradford Hill analysis). It does Defendants no good to note that a biostatistician was once excluded for offering a biological plausibility opinion he did not develop, *see* Defs. Br. at 43 (citing *In re Mirena IUS Levonorgestrel-Related Prod. Liab. Litig. (No. II)*, 341 F. Supp. 3d 213, 246 (S.D.N.Y. 2018)), or that a cardiologist was excluded after he *conceded* he was unqualified to analyze the

studies that formed the basis for his opinions, *id.* (citing *In re Onglyza (Saxagliptin) & Kombiglyze (Saxagliptin & Metformin) Prods. Liab. Litig.*, 93 F.4th 339, 347 (6th Cir. 2024)).

5. Strength of Association

Dr. Ness reliably concluded that the strength of association factor was "partially met." A-7381. She provided examples that the District Court ignored of "other associations in the range of 1.3 to 1.4 that are considered causal," including occupational exposure to benzene and leukemia, estrogen-progestin menopausal hormone therapy (HT) and breast cancer, trichloroethylene and kidney cancer, talcum powder and ovarian cancer, and air pollution and mortality. A-7376–78, A-7237–39.

Defendants claim the District Court could not have erred because of Dr. Ness's conservative conclusions, but the fact that Dr. Ness was conservative on this factor should only confirm that she applied reliable principles, not a "result-driven" methodology.

* * *

The District Court followed its own policy views and crowned the winner of a scientific debate. Defendants contend that Plaintiffs' "attack on the [D]istrict [C]ourt's motivations is inappropriate and unfounded," Defs. Br. at 31, but the unusual lengths to which the District Court went to reach its decisions are simply a

matter of record. Throughout the MDL, the District Court repeatedly referred to the public importance of the litigation. *See, e.g.* 916.A-503 (Tr. 15:16–21); 916.A-562 (Tr. 74:2–13); 916.A-574 (Tr. 86:15–19); 916.A-5369–71 (Tr. 4:21–6:5); 916.A-5537; SPA-179. It ordered Plaintiffs to submit a proposed warning even though they *never* have such a burden, even at trial. 916.A-5721–22 (collecting cases). Plaintiffs' objection on this score was entirely unaddressed, *id.*, and the District Court *sua sponte* solicited FDA's "views" on "Plaintiffs' Proposed warning."[8] MDL Dkt. 561. *None* of these steps have any bearing on whether Plaintiffs' experts reliably applied scientific principles.

What is quite relevant to Rule 702, however, is what independent scientists say and do. The District Court routinely ignored these scientific sources, preferring instead to announce and apply "rules" of science that do not have a shred of support in the peer reviewed literature. The District Court even sprung these rules *sua sponte*, blindsiding Plaintiffs with supposed methodological flaws that not even Defendants attempted to raise. *See. e.g.*, Appeal No. 916 Opening Br. at 52; Appeal No. 916 Reply at 21. *Of course* these unprecedented steps suggest that the District Court's "motivations" were to reach the policy outcome it sincerely believes is

[8] Defendants argue that Plaintiffs did not object to the District Court's unprecedented invitation to FDA, Defs. Br. at 17, but parties cannot object when courts operate on their own initiative, without regard to party presentation.

correct. Defs. Br. at 31. Short of a district court admitting flat out that it is deciding for itself the right answers to scientific questions, it is difficult to imagine a starker departure from Rule 702's gatekeeping role. There is nothing "inappropriate" about bringing these maneuvers to the Court's attention. *Id.*

II. The District Court Erred By Granting Summary Judgment.

Defendants defend two contradictory positions. First, they claim that their *own expert's* prior body of scientific analysis is unreliable. When that fails, they next claim that what he previously said was in fact reliable so long as it is interpreted, well, charitably, by "credit[ing] [his] self-serving explanations or adopt[ing] possible exculpatory interpretations on his behalf." *Palin v. N.Y. Times Co.*, 113 F.4th 245, 264 (2d Cir. 2024). But assessing credibility and interpreting the evidence are tasks reserved for the jury. The Court should reverse.

A. Dr. Faraone's Opinions Are Admissible Under Rule 702.

In a remarkable feat of turnabout, Defendants argue that their own expert's prior scientific work is unreliable under Rule 702. As the Court would expect, Defendants never argued their own expert's work was unreliable below. Their post-hoc attempt to claim they preserved an objection on summary judgment, Defs. Br. at 46–47, does not withstand scrutiny. That "objection" was a single line in their District Court brief claiming Dr. Faraone's testimony was "not comparable to admissible expert testimony in terms of reliability," MDL Dkt. 1506 at 9 (cleaned

up), and double hearsay, *id.* at 10 n.4. "This vague statement is plainly insufficient to preserve the arguments that [Defendants] now make[] on appeal." *United States v. Holley*, 813 F.3d 117, 122 (2d Cir. 2016) (per curiam). The District Court had already *denied* the only Rule 702 motion filed against Dr. Faraone, and Defendants have never challenged that order, because it is the outcome *they* sought. 916.SPA-5.

Forfeiture aside, Defendants claim that Dr. Faraone's opinions are unreliable. But Dr. Faraone's opinions are derived from his peer-reviewed publications and other work product that he created in the normal course of his career as a preeminent scientist. *See* 916.A-4307 (Tr. 329:1–8) (describing his blog posts as part of "this current phase of my career" to encourage people to "engage [] with the evidence"). "The ultimate objective" of the Rule 702 inquiry "is to make certain that an expert . . . employs in the courtroom the same level of intellectual rigor that characterizes the practice of an expert in the relevant field." *Richardson v. Corrective Med. Care Inc.*, No. 22-210, 2023 WL 3490904, at *2 (2d Cir. May 17, 2023) (per curiam) (quoting *Kumho Tire Co.*, 526 U.S. at 152). This objective is easily satisfied where, as here, an expert's opinions are reached in the course of his independent work outside the courtroom. Because Dr. Faraone's opinions are the product of his day-to-day "practice . . . in the relevant field," they are necessarily characterized by "the same level of intellectual rigor" required in his field. *Kumho Tire*, 526 U.S. at 152.

The late-breaking criticisms Defendants lodge against their own expert's statements—i.e., that they are "isolated," or taken "out-of-context," Defs. Br. at 47—"go to the weight, not the admissibility" of his testimony. *Amorgianos*, 303 F.3d at 267 (quotation omitted). Where an expert reaches his opinions in the normal course of his work as a scientist, it can hardly be said that he "lacks 'good grounds' for his . . . conclusion." *Id.* (quoting *In re Paoli R.R. Yard PCB Litig.*, 35 F.3d 717, 746 (3d Cir. 1994)). This is particularly true where the expert's conclusions have been vetted by the peer-review process, as Dr. Faraone's have been here. *See* 916.A-4766 (Faraone (2021)); 916.A-4797 (Khoury (2022)); 916.A-4817 (Joseph 2015)). "Vigorous cross-examination, presentation of contrary evidence, and careful instruction on the burden of proof" are eminently "appropriate means of attacking" Dr. Faraone's pro-causation opinions. *Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 579, 596 (1993) (citation omitted).

B. Dr. Faraone's Opinions Are Sufficient To Establish Causation.

Defendants and the District Court assert that Dr. Faraone did not "state that prenatal exposure to acetaminophen causes ADHD in offspring." Defs. Br. at 48 (quoting SPA-270). That assertion turns summary judgment upside down, resolving all reasonable inferences (and some unreasonable ones, too) in Defendants' favor. Dr. Faraone *expressly* stated that it makes "biological sense[] that [APAP] exposure during pregnancy *could cause* ADHD," 916.A-4307 (Tr. 331:19–332:4) (emphasis

added), and that "*causes of ADHD*" include when "the fetus" is "expose[d] to acetaminophen," 916.A-5347–48 (emphasis added). He gave a slide presentation where the first slide said, "Causes of ADHD" and the very next slide listed prenatal exposure to acetaminophen. A reasonable jury could easily conclude that Dr. Faraone was opining that prenatal APAP exposure causes ADHD.

To take the case from the factfinder, the District Court was determined to draw every conceivable inference in Defendants' favor. *See Palin*, 113 F.4th at 266. In many instances, the District Court's inferences are not even reasonable. For example, Dr. Faraone undeniably said in his prior work that APAP is a risk factor for ADHD. To spin this damning evidence of causation, the District Court simply accepted Dr. Faraone's post-retention explanation that "risk factor" is merely a synonym for "correlate" not "cause." *See* Defs. Br. at 48 (quoting SPA-272). But that dissembling rationale is flatly out of step with how scientists like Dr. Faraone normally use the term "risk factor." The *Dictionary of Epidemiology*—a resource with contributions from over 230 epidemiologists around the world—defines "risk factor" as "[a] factor that is *causally related* to a change in the risk of a relevant health process, outcome, or condition." *A Dictionary of Epidemiology* 251 (Miquel Porta ed., 6th ed. 2016) (emphasis added). "If the relationship is *noncausal* the factor is just a risk marker," which is "a factor that is noncausally associated with the risk of a disease or other specified outcome." *Id.* (emphasis added). A jury should be

permitted to decide whether to believe Dr. Faraone's post hoc justification for his prior statements or the scientific dictionary.

III. Plaintiffs' Claims Are Not Preempted.

Defendants replicate the preemption arguments they made in the earlier appeals, claiming the general pregnancy warning, 21 C.F.R. § 201.63, makes it impossible to comply with state duties to warn. Defs. Br. at 49–56. But nothing in § 201.63 prevented Defendants from adding a warning about ADHD. The regulation simply requires a general pregnancy warning, and 21 C.F.R. § 330.1(c)(2) merely requires that it be given with the "exact language" provided. Neither regulation precludes additional warnings, as the actual APAP labels confirm. There is no FDA interpretation to the contrary, and even if there were, it would not be entitled to *Auer* deference.

A. The General Pregnancy Warning Does Not Preempt Plaintiffs' Claims.[9]

The plain text of § 201.63 requires a general pregnancy warning. It says nothing to forbid *additional* warnings. The provision states:

> The labeling for all over-the-counter (OTC) drug products that are intended for systemic absorption, unless specifically exempted, shall contain a general warning

[9] APAP is subject to the Internal Analgesic, Antipyretic, and Antirheumatic Drug Products for Over-the-Counter Human Use; Tentative Final Monograph, 53 Fed. Reg. 46,204, 46,255 (Nov. 16, 1988) ("Acetaminophen Monograph"), as well as the general federal regulations for OTC drug labeling, including the general pregnancy warning found in 21 C.F.R. § 201.63(a).

> under the heading "Warning" (or "Warnings" if it appears with additional warning statements) as follows: "If pregnant or breast-feeding, ask a health professional before use." [first four words of this statement in bold type] In addition to the written warning, a symbol that conveys the intent of the warning may be used in labeling.

21 C.F.R. § 201.63. There is not a shred of regulatory text that forbids a manufacturer from using the required warning *and also* adding warnings required by state law. The plain text of the regulations is dispositive in their sweep. *See Isett v. Aetna Life Ins.*, 947 F.3d 122, 132 (2d Cir. 2020) (declining to read in limitations not included in a regulation's express terms). Defendants cannot meet their "heavy burden" to show that it is impossible to satisfy both state and federal law. *Wyeth v. Levine*, 555 U.S. 555, 567–69 (2009).[10]

B. 21 C.F.R. § 330.1(c)(2) Does Not Prohibit Additional Pregnancy Warnings.

Defendants' invocation of the "exact language" rule, 21 C.F.R. § 330.1(c)(2), is a non sequitur. Defs. Br. at 51. That rule merely requires marketers to state the general pregnancy warning verbatim. It says nothing about whether a marketer is forbidden from making an *additional* warning. In fact, the exact-language rule makes clear that it is limited to warnings "established and identified by quotation

[10] The District Court agreed: "The Pregnancy Warning Regulation simply does not speak to whether a further warning related to a drug's use during pregnancy can be added to the general Pregnancy Warning on a drug label . . . ." 916.A-965; *see also* 916.A-599–600 ("[Section 330.1] does not address the ability of manufacturers to supplement the general warning with safety warnings specific to their OTC drug.").

marks in an applicable OTC drug monograph or by regulation." 21 C.F.R. § 330.1(c)(2). Where a warning has not been "established" by a regulation or monograph—which the ADHD warning has not been—the marketer is free to add it with truthful, relevant language required by state law.

Defendants' own APAP labels confirm this textual point. The Tylenol label warns: "Do not use . . . if you are allergic to acetaminophen," *see* MDL Dkt. 774 ¶ 111, but the Acetaminophen Monograph does not require this warning. *See* Acetaminophen Monograph, 53 Fed. Reg. at 46,256–57. The Tylenol label also warns that "acetaminophen may cause severe skin reactions" such as a "rash." MDL Dkt. 774 ¶ 111. This warning is not mentioned anywhere in the monograph either, though FDA has *encouraged* marketers to *voluntarily* add it.[11] In short, if the "exact language" rule somehow limited marketers to warnings contained in regulations or the applicable monograph, the Tylenol label would be—and has long been—in violation of the law. *See* Defs. Br. at 51.

Defendants respond with a "pregnancy is special" rule of law, suggesting that the exact-language rule makes the general pregnancy warning both a floor and a ceiling on the subject of pregnancies, while marketers are free to add additional warnings on other conditions. *See, e.g.*, Defs. Br. at 51. This argument of

[11] FDA, *Guidance for Industry: Recommended Warning for Over-the-Counter Acetaminophen-Containing Drug Products and Labeling Statements Regarding Serious Skin Reactions* (Jan. 2017), https://www.fda.gov/media/90572/download.

convenience ignores that the clear terms of § 330.1(c)(2) apply to the entire Acetaminophen Monograph *and* applicable regulations. *See* 21 C.F.R. § 330.1(c)(2) (requiring exact language when established "by quotation marks *in an applicable OTC drug monograph or by regulation (e.g., § 201.63 of this chapter)*" (emphasis added)). Either this regulation makes the entire Acetaminophen Monograph and the general pregnancy warning exclusive or it does not; there is no special carveout for a certain type of warning.

Thwarted by the text, Defendants point to regulatory history. Defs. Br. at 53–54. But the cited informal FDA statements about exclusivity describe an *old* regulatory regime—before § 330.1(c)(2) had even been proposed—and necessarily lack any consideration of the yet-to-be-written rule. *See* Labeling of Drug Products for Over-the-Counter Human Use, 50 Fed. Reg. 15,810, 15,810 (Apr. 22, 1985) ("proposing to change [FDA's] 'exclusivity' policy" by promulgating § 330.1(c)(2)); 916.A-967 (The District Court stating that Defendants' same argument "misconstrues both the text of the Exact Language Regulation and the history behind its creation.").

To the extent FDA's informal statements are relevant, Defendants ignore FDA's other statements that reinforce the plain meaning of § 330.1(c)(2). As the District Court recognized, FDA later finalized § 201.66 and OTC drug labeling requirements and changed its view on where the additional warnings should appear.

See 916.A-590–91; 916.A-954–55. In response to comments where "several manufacturers requested that FDA allow voluntary warnings to appear under the appropriate headings to further protect consumers from possible misuse of the product," FDA encouraged "manufacturers to discuss with the agency the addition of *voluntary warnings* to OTC drug products." Over-the-Counter Human Drugs; Labeling Requirements, 64 Fed. Reg. 13,254, 13,271 (Mar. 17, 1999) (emphasis added); *see* 916.A-590–91 (same); 916.A-954–55, 916.A-967.

C. FDA Did Not Interpret § 201.63 To Forbid Additional Pregnancy Warnings, And *Auer* Deference Does Not Apply.

Defendants urge the Court to defer to FDA's interpretation of 21 C.F.R. § 201.63 as "exclusive and preemptive," but FDA did not interpret the text. Defs. Br. at 54. FDA said, rather, in 1982, that "differing State OTC drug pregnancy-nursing warning requirements would prevent accomplishment of the *full purpose and objectives* of the agency in issuing the regulation." Pregnant or Nursing Women; Delegations of Authority and Organization; Amendment of Labeling Requirements for Over-the-Counter Human Drugs, 47 Fed. Reg. 54,750, 54,756 (Dec. 3, 1982) (emphasis added). That out-of-date commentary is about objects-and-purposes preemption, not impossibility, which is all Defendants argued below. It does not purport to interpret a single ambiguous word of the general pregnancy warning. *See Kisor v. Wilkie*, 588 U.S. 558, 563 (2019) (stating that *Auer* deference is about deferring to "agencies' reasonable readings of [their own] genuinely

ambiguous regulations"); *Gonzales v. Oregon*, 546 U.S. 243, 257 (2006) (an agency's comments "cannot be considered an interpretation of the regulation" unless the agency is interpreting "the meaning of the regulation"). FDA's commentary was also focused on "whether States can modify the Pregnancy Warning or replace it with their own version," not whether an additional warning could be added. MDL Dkt. 601 at 4.

Defendants also never engage with the rigorous steps *Kisor* insists upon before yielding to an agency's view. Defendants have not shown that (1) § 201.63 is ambiguous; (2) it would be reasonable to read into § 201.63 a different (and unwritten) command; (3) FDA's desire to preempt the laws of sovereign states somehow falls within its substantive expertise, *see Wyeth*, 555 U.S. at 576–77 ("[A]gencies have no special authority to pronounce on pre-emption" and thus courts have "not deferred to an agency's *conclusion* that state law is pre-empted."); and (4) FDA's 1982 interpretation was "consistent" with later statements (it was not) and "contemporaneous" with a regulation promulgated in 1999 (it was not). *Kisor*, 588 U.S. at 577-79 (holding that these elements are required for deference).

In sum, nothing in federal law precludes Plaintiffs' claims.

CONCLUSION

The judgment of the District Court should be reversed.

CERTIFICATE OF COMPLIANCE

This document complies with the type-volume limit of Fed. R. App. P. 32(a)(7)(B) and Local Rule 32.1(a)(4)(B) because, excluding the parts of the document exempted by Fed. R. App. P. 32(f), this document contains 6,883 words.

This document complies with the typeface requirements of Fed. R. App. P. 32(a)(5) and the type-style requirements of Fed. R. App. P. 32(a)(6) because this document has been prepared in a proportionally spaced typeface using Microsoft 365 Word in Times New Roman 14-pt font.

Dated this 11th Day of April, 2025.

/s/ Ashley Keller
Ashley Keller
KELLER POSTMAN LLC
Attorneys for Plaintiffs–Appellants

CERTIFICATE OF SERVICE

I hereby certify that, on April 11, 2025, an electronic copy of the foregoing Reply Brief was filed with the Clerk of Court using the ECF system and thereby served upon all counsel appearing in this case.

/s/ Ashley Keller
Ashley Keller
KELLER POSTMAN LLC
Attorneys for Plaintiffs–Appellants